BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e112053
doi: 10.3897/biss.7.112053

OPEN ACCESS

Conference Abstract

# Ten Years and a Million Links: Building a global taxonomic library connecting persistent identifiers for names (LSIDs), publications (DOIs), and people (ORCIDs)

Roderic Page ‡

‡ University of Glasgow, Glasgow, United Kingdom

Corresponding author: Roderic Page (roderic.page@glasgow.ac.uk)

Received: 02 Sep 2023 | Published: 05 Sep 2023

Citation: Page R (2023) Ten Years and a Million Links: Building a global taxonomic library connecting persistent identifiers for names (LSIDs), publications (DOIs), and people (ORCIDs). Biodiversity Information Science and Standards 7: e112053. https://doi.org/10.3897/biss.7.112053

## Abstract

One thing the field of biodiversity informatics has been very good at is creating databases. However, this success in creation has not been matched by equivalent success in creating deep links between records in those databases. Instead, we create an ever growing number of silos. An obvious route to "silo-breaking" is the shared use of the same persistent identifiers for the same entities across those databases. For example, we have minted millions of Life Science Identifiers (LSIDs) for taxonomic names (which can be resolved at lsid.io), and a growing number of taxonomic papers have Digital Object Identifiers (DOIs), but we lack connections between these two identifiers. In this talk I describe work over the last decade to make these connections between LSIDs and DOIs across three large taxonomic databases: Index Fungorum, International Plant Names Index (IPNI), and the Index to Organism Names (ION) (Page 2023). Over a million names have been matched to DOIs or other persistent identifiers for taxonomic publications (Fig. 1 shows the coverage of publications for animal names). This represents approximately 36% of animal, plant or fungal names for which publication data is available.

© Page R. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

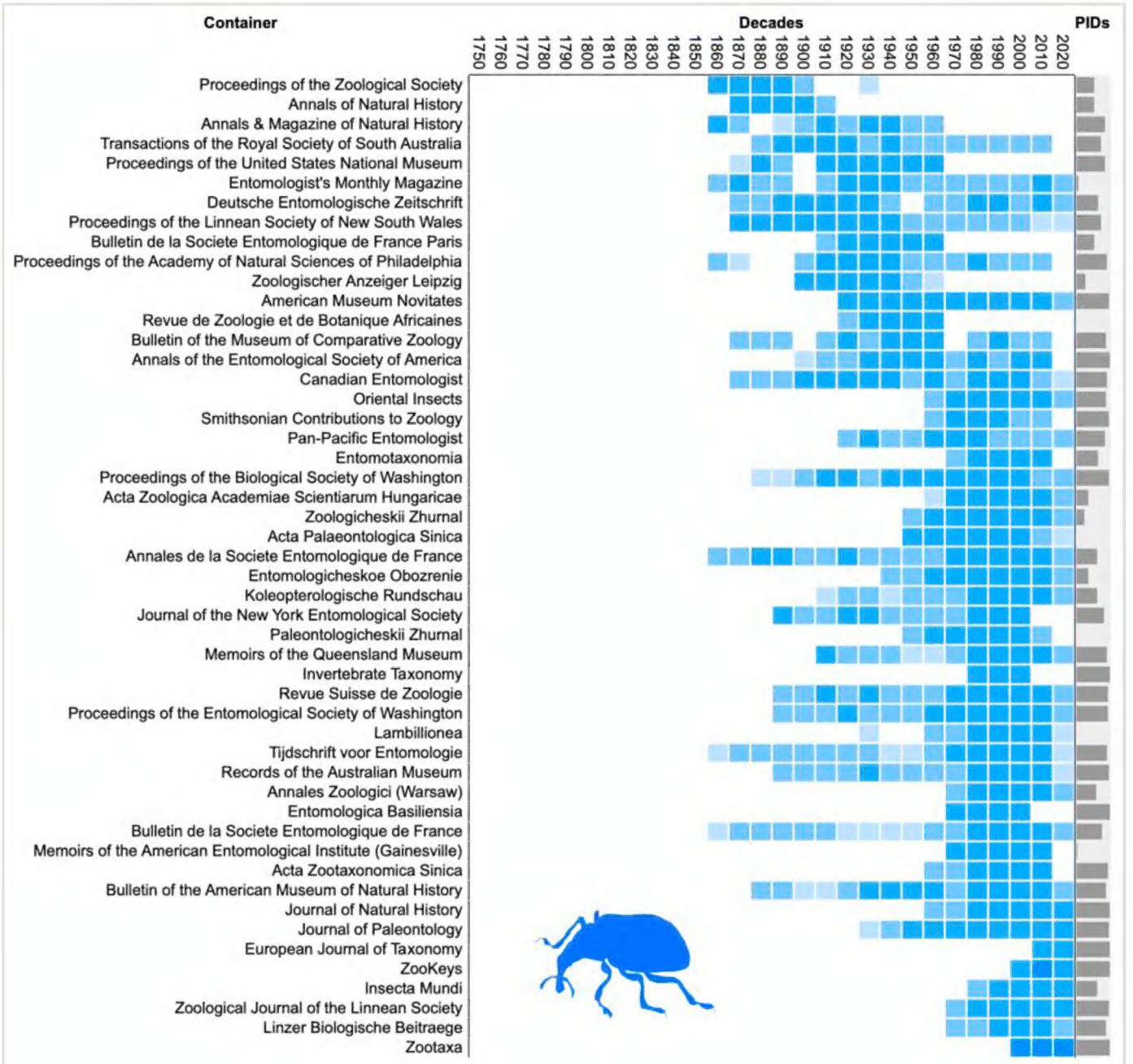


Figure 1.

Density distribution of taxonomic names published in the decades from 1750 to 2020 across the 50 publication venues (“containers”) that published the most names for animals. The intensity of colour in each of square in the diagram is proportional to the number of names published in that decade for each publication. The containers are ordered by the decade in which they published most names. The column labelled “PIDs” shows bars proportional to the percentage of names in each publication that have been linked to a persistent identifier for that publication. Figure from Page (2023).

The mappings between LSIDs and publication persistent identifiers (PIDs) such as DOIs and Wikidata item identifiers, are made available through ChecklistBank (datasets 129659 164203, 128415), and also archived in Zenodo. By combining these LSID and DOI links with Open Researcher and Contributor ID (ORCIDs) for taxonomists, we can potentially gain insight into who is doing taxonomic research, where they work, and how they are funded. Possible applications of this data are discussed, including a tool to discover the citation for a species name (Species Cite, Fig. 2), using DOI to ORCIDs to discover who is doing taxonomic research, and creating a linked data version of the Catalogue of Life.

Figure 2.

Screenshot of Species Cite showing details for a taxonomic name, with a formatted citation, links to persistent identifiers (e.g., DOI, Wikidata), and a PDF of the publication. Figure by the author.

## Keywords

linked data, biodiversity knowledge graph, Index Fungorum, IPNI, ION, Catalogue of Life

## Presenting author

Roderic Page

## Presented at

TDWG 2023

## Conflicts of interest

The authors have declared that no competing interests exist.

## References


- Page RM (2023) Ten years and a million links: building a global taxonomic library connecting persistent identifiers for names, publications and people. bioRxiv https://doi.org/10.1101/2023.05.29.542697